

## Robust TDT-type candidate-gene association tests

G. ZHENG<sup>1</sup>, B. FREIDLIN<sup>2</sup> AND J. L. GASTWIRTH<sup>2,3</sup>

<sup>1</sup>National Heart, Lung and Blood Institute, Bethesda, MD

<sup>2</sup>National Cancer Institute, Bethesda, MD

<sup>3</sup>George Washington University, Washington, DC

(Received 18.7.01. Accepted 13.12.01)

### SUMMARY

In studies of association between genetic markers and a disease, the transmission disequilibrium test (TDT) has become a standard procedure. It was introduced originally as a test for linkage in the presence of association and can be used as a test for association under appropriate assumptions. The power of the TDT test for association between a candidate gene and disease depends on the underlying genetic model and the TDT is the optimal test if the additive model holds. Related methods have been obtained for a given mode of inheritance (e.g. dominant or recessive). Quite often, however, the true model is unknown and selection of a single method of analysis is problematic, since use of a test optimal for one genetic model usually leads to a substantial loss of power if another genetic model is the true one. The general approach of efficiency robustness has suggested two types of robust procedures, which we apply to TDT-type association tests. When the plausible range of alternative models is wide (e.g. dominant through recessive) our results indicate that the maximum (MAX) of several test statistics, each of which is optimal for quite different models, has good power under all genetic models. In situations where the set of possible models can be narrowed (e.g. dominant through additive) a simple linear combination also performs well. In general, the MAX has better power properties than the TDT for the study of candidate genes when the mode of inheritance is unknown.

### INTRODUCTION

The transmission disequilibrium test (TDT) described in Spielman *et al.* (1993) was developed as a test of linkage between a marker locus and a disease in the presence of allelic association, and can also be used in testing for association between a candidate gene and a disease. The TDT followed earlier work (Rubinstein *et al.* 1981; Field *et al.* 1986; Falk & Rubinstein, 1987; Ott, 1989; Thomson *et al.* 1989; Terwilliger & Ott, 1992) on identification of disease loci by combining both the linkage and population association approaches. Population substructure does not affect the use of the TDT as a test for linkage. The TDT is valid as a test for association when simplex families are studied, however, under certain circumstances it will detect association due to population substructure (see scenario 3 in Ewens & Spielman, 1995). For a further discussion of these issues see Spielman & Ewens (1996) and Curnow, Morris & Whittaker (1998).

By focusing on a specific genetic model in the candidate gene setting more powerful tests can be obtained. Schaid & Sommer (1994) derived TDT-type statistics that are, respectively, most powerful for dominant and recessive models, and showed that the TDT is optimal under an additive or multiplicative model. Since the underlying genetic model is usually not known in advance, there is

Correspondence: Gang Zheng, PhD Office of Biostatistics Research, National Heart, Blood and Lung Institute, 6701 Rockledge Drive, MSC 7938, Bethesda, MD 20892, USA. Tel: (301) 435 1287; Fax: (301) 480 1862.

E-mail: zhengg@nhlbi.nih.gov

a need to develop a test that has high power over the set of plausible genetic models. Results from efficiency robustness (Gastwirth, 1966; Gastwirth & Freidlin, 2000; Shih & Whittemore, 2001) suggest two possible procedures. When the underlying genetic model can range from a recessive to a dominant one, the maximum of several test statistics is shown to have good power under all genetic models. When the set of plausible models is smaller, e.g. can range from additive to dominant, a simple linear combination performs well.

#### MODELS AND TESTS

Schaid & Sommer (1993, 1994) developed a likelihood approach to testing for association between a candidate gene and a disease. One affected child per family and both parents are sampled and genotyped. By conditioning on parental genotypes they obtained tests that do not require the assumption of Hardy-Weinberg equilibrium (HWE). In the candidate gene setting, conditioning on the parental genotypes implies that the contributions from the two parents are independent under the null hypothesis of no association between the disease and candidate gene. Denote the candidate disease allele by  $D$ , its complement (normal allele) by  $d$  and the population frequency of  $D$  by  $p$ . There are 6 different mating types (1)  $DD \times DD$ , (2)  $DD \times Dd$ , (3)  $DD \times dd$ , (4)  $Dd \times Dd$ , (5)  $Dd \times dd$ , and (6)  $dd \times dd$ . Let  $f_0 = P(DS|dd)$ ,  $f_1 = P(DS|Dd)$  and  $f_2 = P(DS|DD)$  denote the penetrances, where  $DS$  is the event that an individual has the disease. In terms of relative risks  $r_1 = f_1/f_0$  and  $r_2 = f_2/f_0$  the four basic genetic models are (1) dominant ( $D$ ):  $r_1 = r$ ,  $r_2 = r$ , (2) recessive ( $R$ ):  $r_1 = 1$ ,  $r_2 = r$ , (3) additive ( $A$ ):  $r_1 = r$ ,  $r_2 = 2r - 1$ , (4) multiplicative ( $M$ ):  $r_1 = r$ ,  $r_2 = r^2$ . Conditional on the parental mating type, the distributions of the case genotype do not require HWE to hold. For mating types 1, 3 and 6 these conditional distributions are degenerate, with only one offspring genotype possible: these matings are thus not informative. For mating type 2 the two possible disease case genotypes,  $DD$  and  $Dd$ , have binomial distribution with parameter  $P(\text{case} = DD | \text{mating type} = 2) = r_2/(r_1 + r_2)$ . For mating type 4 the three possible case genotypes,  $DD$ ,  $Dd$  and  $dd$ , have a trinomial distribution with parameters  $P(\text{case} = DD | \text{mating type} = 4) = r_2/(2r_1 + r_2 + 1)$  and  $P(\text{case} = Dd | \text{mating type} = 4) = 2r_1/(2r_1 + r_2 + 1)$ . For mating type 5 the two possible disease case genotypes,  $Dd$  and  $dd$ , have binomial distribution with parameter  $P(\text{case} = Dd | \text{mating type} = 5) = r_1/(r_1 + 1)$ . Each of the models is parameterized in terms of the single parameter  $r$ , specifying the potential increased risk of disease. Thus the null hypothesis of no association between the disease and the candidate gene, i.e.  $f_2 = f_1 = f_0$ , reduces to testing  $H_0: r = 1$  vs.  $H_1: r > 1$ . Note again that in certain circumstances, a positive association may be due to population subdivision. An efficient score statistic (Rao, 1973) can be used to obtain an optimal test. For the four basic models  $D$ ,  $R$ ,  $M$  and  $A$  (Table 1), the statistics are (see appendix A):

$$Z_D = \frac{(n_{42} + n_{41} - 3n_{44}/4) + (n_{51} - n_{50}/2)}{\sqrt{3n_{44}/16 + n_{50}/4}} \quad (1)$$

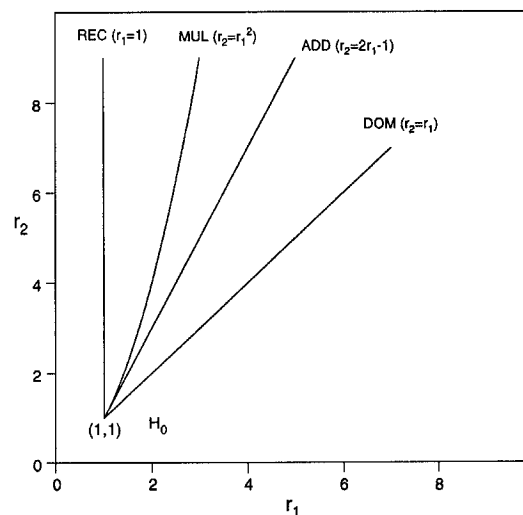
$$Z_R = \frac{(n_{22} - n_{21}/2) + (n_{42} - n_{44}/4)}{\sqrt{n_{22}/4 + 3n_{44}/16}} \quad (2)$$

$$Z_A = Z_M = \frac{(n_{22} - n_{21}) + 2(n_{42} - n_{44}) + (n_{51} - n_{50})}{\sqrt{n_{22} + 2n_{44} + n_{50}}} \quad (3)$$

where  $n_i$  is the total number of affected from parental mating type  $i$  and  $n_{ij}$  is the number of affected from parental mating type  $i$  that have  $j$   $D$  alleles. Note that test statistics  $Z_A$  and  $Z_M$  are equivalent to the TDT. Efficient score tests attain their optimal properties as the parameter of interest converges to its null value, i.e.  $r \rightarrow 1$ . As the first order Taylor series approximation of  $r^{1/2}$  near

Table 1. Probabilities of case genotypes for the four models given mating types

Mating type	Case Genotype	Counts	Conditional probability			
			D	R	A	M
2: $DD \times Dd$	DD	$n_{22}$	1/2	$r/(r+1)$	$(2r-1)/(3r-1)$	$r/(r+1)$
	Dd	$n_{21}$	1/2	$1/(r+1)$	$r/(3r-1)$	$1/(r+1)$
4: $Dd \times Dd$	DD	$n_{42}$	$r/(3r+1)$	$r/(r+3)$	$(2r-1)/4r$	$r^2/(r+1)^2$
	Dd	$n_{41}$	$2r/(3r+1)$	$2/(r+3)$	1/2	$2r/(r+1)^2$
	dd	$n_{40}$	$1/(3r+1)$	$1/(r+3)$	$1/4r$	$1/(r+1)^2$
5: $Dd \times dd$	Dd	$n_{51}$	$r/(r+1)$	1/2	$r/(r+1)$	$r/(r+1)$
	dd	$n_{50}$	$1/(r+1)$	1/2	$1/(r+1)$	$1/(r+1)$

Fig. 1. Four genetic models in the  $(r_1, r_2)$  plane.

$r = 1$  is  $(1+r)/2$  the additive and multiplicative models are essentially the same (see also Figure 1 and the accompanying discussion). This explains the equivalence of  $Z_A$  and  $Z_M$ .

In most situations, however, the true underlying model is unknown and selection of a single method of analysis is problematic since use of any one optimal test may lead to a loss of power under another model. A test is efficiency robust over a class of models if its relative efficiency compared to the optimal test for each of the models is high when the data come from that model. As the efficiency of a test is defined in terms of its power to detect a local alternative, such procedures typically have good power properties over the range of the possible models. The Maximin Efficiency Robust Test (MERT) has the highest minimum efficiency relative to the optimal tests (Gastwirth, 1966, 1985) and was used in Gastwirth & Freidlin (2000) to examine tests based on affected sib-pairs and triples proposed by Whittemore & Tu (1998). The MERT is obtained as a linear combination of the various optimal test statistics given above. In many cases it is equal to the standardized sum of the test statistics optimal for the two most extreme models (extreme pair). Alternatively, the maximum of several optimal test statistics (MAX) is sometimes employed; this approach has been shown to have attractive power properties in a variety of applications (e.g. Tarone, 1981; Schaid & Nick, 1990; Fleming & Harrington, 1991). Efficiency robust principles suggest that the test statistics for the extreme models be used. Furthermore, when their null-correlation is low the optimal test statistic for an intermediate model should be included. While the distributional properties of the MAX are more complicated than those of asymptotically normal test statistics, wide availability of powerful

software has resulted in their increased popularity in recent years. Both methods depend on the null correlation of the optimal statistics of their components. A short review of these robust methods is provided in appendix B.

First, consider a family of four models  $D$ ,  $R$ ,  $A$  and  $M$ . The null correlation matrix of the test statistics  $Z_D$ ,  $Z_R$  and  $Z_A$  when the null hypothesis of no association, i.e.  $H_0: r = 1$ , (i.e.,  $f_2 = f_1 = f_0$ ) is (details available upon request):

$$\begin{array}{ccc} Z_D & Z_R & Z_A \\ 1 & \frac{n_4}{\sqrt{3n_4+4n_5}\sqrt{3n_4+4n_2}} & \frac{2n_4+2n_5}{\sqrt{3n_4+4n_5}\sqrt{n_2+2n_4+n_5}} \\ & 1 & \frac{2n_4+2n_2}{\sqrt{3n_4+4n_2}\sqrt{n_2+2n_4+n_5}} \\ & & 1 \end{array} \quad (4)$$

From the correlation structure it is easy to show that the correlation between the test statistics  $Z_D$  and  $Z_R$  is the smallest and has an upper bound of  $1/3$ . Figure 1 gives a graphical representation of the four genetic models in the  $(r_1, r_2)$  plane. Dominant, recessive and additive models are represented by corresponding rays and the multiplicative model by a quadratic curve with origin at  $(1,1)$ , the null hypothesis. All other genetic models for various values of  $r_1$  and  $r_2$  are represented by rays with the same origin lying between the dominant and recessive models. Using the results of Gastwirth (1985) it can be shown that the MERT for the entire family is the MERT for the extreme pair, the two least correlated test statistics in the family,  $Z_D$  and  $Z_R$  (results are available on request):

$$\text{MERT} = \frac{Z_D + Z_R}{\sqrt{2(1 + \text{corr}(Z_D, Z_R))}} \quad (5)$$

Furthermore, it can be shown that the addition of the multiplicative model to the family does not change the MERT (details are available on request).

From (4) and (5) one can show that the correlation between the MERT statistic and the TDT statistic is given by

$$\text{corr}(\text{MERT}, \text{TDT}) = \frac{\text{corr}(Z_D, \text{TDT}) + \text{corr}(Z_R, \text{TDT})}{\sqrt{2(1 + \text{corr}(Z_D, Z_R))}}.$$

This correlation depends on the proportions of the various parental mating types, which in turn depend on the frequency ( $p$ ) of the candidate allele. Assuming HWE, the correlation between the MERT statistic and the TDT statistic ranges from 0.73 when  $p = 0.001$  to 0.95 when  $p = 0.20$ .

The minimum correlation,  $\rho^*$ , of the optimal statistics for the potential models reflects how broad the family is. When the family of possible genetic models is large, e.g. ranging from recessive to dominant, this minimum correlation is low, so prior results of Freidlin, Podgor & Gastwirth (1999), reviewed in Shih & Whittemore (2001), indicate that the MAX should have higher power across the range of models than the MERT.

#### SIMULATION STUDY

To evaluate the MERT and MAX tests we conducted simulations under various scenarios corresponding to situations where allele frequency  $p$  is constant across the families (HWE holds), as well as situations where the assumption of random mating is violated, i.e.  $p$  varies across population subgroups. First, we generated the genotypes of the parents and one offspring. If the child was affected, the trio was included in the sample. The process was repeated until the stated sample size was obtained.

Table 2(a). Empirical power estimates: HWE holds,  $p = 0.2$  (sample size 100, 5000 replications)

Under model	$Z_D$	$Z_R$	$Z_A$ (TDT)	Test					
				MERT (D, R, A)	MAX (D, R, A)	MERT (D, A)	MAX (D, A)	MERT (R, A)	MAX (R, A)
$H_0$	0.049	0.047	0.049	0.050	0.046	0.049	0.050	0.048	0.051
$D^1$	0.812	0.068	0.720	0.535	0.734	0.790	0.790	0.356	0.625
$R^2$	0.067	0.862	0.492	0.663	0.797	0.233	0.429	0.787	0.821
$A^3$	0.778	0.258	0.811	0.734	0.756	0.820	0.812	0.613	0.732
$M^4$	0.669	0.433	0.799	0.775	0.743	0.762	0.778	0.715	0.749

<sup>1</sup>Model Dominant  $f_0 = 0.02$   $f_1 = 0.045$   $f_2 = 0.045$ .<sup>2</sup>Model Recessive  $f_0 = 0.02$   $f_1 = 0.02$   $f_2 = 0.077$ .<sup>3</sup>Model Additive  $f_0 = 0.02$   $f_1 = 0.0425$   $f_2 = 0.065$ .<sup>4</sup>Model Multiplicative  $f_0 = 0.02$   $f_1 = 0.038$   $f_2 = 0.0722$ .Table 2(b).  $p$ -values of  $Z_A$  (TDT) vs. MAX(D, R, A) (5000 replications)

Under dominant		MAX			
$Z_A$ (TDT)		< 0.01	0.01-0.05	0.05-0.1	> 0.1
< 0.01	2086	245	0	0	
0.01-0.05	380	652	215	23	
0.05-0.1	21	189	122	169	
> 0.1	9	88	121	680	
Recessive		MAX			
$Z_A$ (TDT)		< 0.01	0.01-0.05	0.05-0.1	> 0.1
< 0.01	1232	74	0	0	
0.01-0.05	860	226	62	6	
0.05-0.1	332	132	51	56	
> 0.1	648	478	241	602	
Additive		MAX			
$Z_A$ (TDT)		< 0.01	0.01-0.05	0.05-0.1	> 0.1
< 0.01	2542	400	0	0	
0.01-0.05	133	608	331	40	
0.05-0.1	7	61	93	201	
> 0.1	4	24	49	507	
Multiplicative		MAX			
$Z_A$ (TDT)		< 0.01	0.01-0.05	0.05-0.1	> 0.1
< 0.01	2406	484	0	0	
0.01-0.05	110	633	327	35	
0.05-0.1	9	46	77	239	
> 0.1	2	26	51	555	

Tables 2(a)–4(a) present empirical power estimates of the three optimal test statistics,  $Z_D$ ,  $Z_R$ ,  $Z_A$ , the MERT and MAX for the entire family, as well as MERT and MAX tests for the families where one of the extreme models ( $D$  or  $R$ ) can be ruled out. The simulations are for 0.05 level tests under alternatives chosen so that the optimal tests have approximately 80% power.

When the family of the possible models is wide, e.g. the entire  $R$ – $D$  range, the minimum correlation of the optimal test statistics is low (about 0.12 when HWE holds and  $p = 0.2$ ), so we expect the MAX to be the more robust test. An examination of Tables 2(a)–4(a) confirms this. Further insight into the advantage of MAX over  $Z_A$  (TDT) can be gained from tabulating the joint frequencies of their  $p$ -values calculated on the same data sets. These are reported for < 0.01, 0.01–0.05, 0.05–0.1, > 0.1 categories in Tables 2(b)–4(b). The numbers in the lower left triangle indicate that the MAX attains a lower  $p$ -value than  $Z_A$  while the reverse is reflected in the upper right triangle. For the recessive model, the MAX provides a considerable gain in power over the TDT while losing relatively little power when the additive model holds. Indeed, Table 2(b) indicates that for the recessive model the

Table 3(a). Empirical power estimates: HWE does not hold an equal mixture of populations with  $p = 0.2$  and  $p = 0.01$  (sample size 100, 5000 replications)

Under model	$Z_D$	$Z_R$	$Z_A$ (TDT)	Test					
				MERT (D, R, A)	MAX (D, R, A)	MERT (D, A)	MAX (D, A)	MERT (R, A)	MAX (R, A)
$H_0$	0.051	0.044	0.052	0.047	0.049	0.052	0.053	0.047	0.044
$D^1$	0.800	0.065	0.703	0.509	0.719	0.778	0.777	0.330	0.613
$R^2$	0.057	0.813	0.448	0.605	0.738	0.217	0.391	0.737	0.765
$A^3$	0.782	0.229	0.801	0.717	0.764	0.821	0.814	0.600	0.732
$M^4$	0.587	0.372	0.726	0.699	0.662	0.692	0.698	0.635	0.669

<sup>1</sup>Model Dominant  $f_0 = 0.02$   $f_1 = 0.053$   $f_2 = 0.053$ <sup>2</sup>Model Recessive  $f_0 = 0.02$   $f_1 = 0.02$   $f_2 = 0.1$ <sup>3</sup>Model Additive  $f_0 = 0.02$   $f_1 = 0.051$   $f_2 = 0.082$ <sup>4</sup>Model Multiplicative  $f_0 = 0.02$   $f_1 = 0.042$   $f_2 = 0.0882$ Table 3(b).  $p$ -values of  $Z_A$  (TDT) vs. MAX(D, R, A) (5000 replications)

Under dominant	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	1962	243	0
	0.01-0.05	353	733	211
	0.05-0.1	21	190	147
	> 0.1	4	88	116
Recessive	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	1107	79	0
	0.01-0.05	697	260	87
	0.05-0.1	287	135	56
	> 0.1	536	588	289
Additive	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	2431	428	0
	0.01-0.05	143	696	280
	0.05-0.1	6	75	92
	> 0.1	5	37	60
Multiplicative	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	1909	492	0
	0.01-0.05	105	688	410
	0.05-0.1	10	61	83
	> 0.1	10	33	55

$p$ -value of MAX is  $< 0.01$  while the  $p$ -value of  $Z_A$  (TDT) is  $> 0.10$  in 13% of the simulations. On the other hand, under the  $A$  and  $M$  models, for which  $Z_A$  is optimal, we did not observe any cases where the  $p$ -value of  $Z_A$  was  $< 0.01$  and that of MAX  $> 0.1$ . The results of Tables 3 and 4, which do not assume HWE, indicate that MAX remains the most robust test.

In situations where we can restrict the models so that either the dominant or recessive mode of inheritance can be eliminated on scientific grounds, the MAX of  $Z_A$  and  $Z_R$  (or  $Z_D$ ) remains the most robust. However, the MERT or  $Z_A$  are relatively efficiency robust and may be easier to apply as they are asymptotically normal.

While Tables 2-4 cover simulations performed under the 4 alternative genetic models, we also conducted a series of simulations assuming several intermediate models. The results were similar to ones discussed and are not presented.

Table 4(a). Empirical power estimates: HWE does not hold a 1:9 mixture of populations with  $p = 0.3$  and  $p = 0.01$  (sample size 100, 5000 replications)

Under model	$Z_D$	$Z_R$	$Z_A$ (TDT)	Test					
				MERT (D, R, A)	MAX (D, R, A)	MERT (D, A)	MAX (D, A)	MERT (R, A)	MAX (R, A)
$H_0$	0.053	0.025	0.047	0.046	0.045	0.0452	0.051	0.043	0.040
$D^1$	0.806	0.063	0.672	0.486	0.711	0.770	0.782	0.30	0.556
$R^2$	0.055	0.793	0.522	0.582	0.709	0.252	0.441	0.727	0.742
$A^3$	0.773	0.229	0.792	0.695	0.746	0.817	0.806	0.552	0.698
$M^4$	0.574	0.502	0.791	0.774	0.722	0.733	0.751	0.724	0.737

<sup>1</sup>Model Dominant  $f_0 = 0.02$   $f_1 = 0.085$   $f_2 = 0.085$ <sup>2</sup>Model Recessive  $f_0 = 0.02$   $f_1 = 0.02$   $f_2 = 0.169$ <sup>3</sup>Model Additive  $f_0 = 0.02$   $f_1 = 0.079$   $f_2 = 0.138$ <sup>4</sup>Model Multiplicative  $f_0 = 0.02$   $f_1 = 0.06$   $f_2 = 0.18$ Table 4(b).  $p$ -values of  $Z_A$  (TDT) vs. MAX(D, R, A) (5000 replications)

Under dominant	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	1711	266	0
	0.01-0.05	417	725	215
	0.05-0.1	47	212	130
	> 0.1	14	162	188
Recessive	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	1331	103	0
	0.01-0.05	617	430	103
	0.05-0.1	179	195	75
	> 0.1	227	460	263
Additive	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	2204	446	0
	0.01-0.05	201	720	340
	0.05-0.1	14	93	85
	> 0.1	3	48	68
Multiplicative	MAX			
	$Z_A$ (TDT)	< 0.01	0.01-0.05	0.05-0.1 > 0.1
	< 0.01	2195	562	0
	0.01-0.05	100	651	386
	0.05-0.1	4	62	55
	> 0.1	7	27	37

In designing a study it is important to remember that only mating types 2, 4 and 5 are informative for association. Moreover, for small  $p$  ( $\leq 0.01$ ) the expected fraction of the informative sample that comes from mating types 2 and 4 is quite low. As  $Z_R$  is based on only those two mating types, the number of cases actually contributing to  $Z_R$  is very small. This may explain the inaccuracy of the normal approximation to the size of  $Z_R$  in Table 4(a).

## DISCUSSION

An alternative to using a robust test is to employ an adaptive procedure that uses the observed data to identify the model and select the appropriate test. In this setting, an adaptive procedure could only produce a marginal gain in power since the power of the MAX ( $Z_R, Z_A, Z_D$ ) is quite close to the power of the optimal test.

Another approach, reviewed by Slager *et al.* (2001), uses likelihood ratio tests for different models and then chooses the lowest  $p$ -value, corrected for multiple comparisons; this method is similar to a proposal of Sham (1998). Using MAX enables one to obtain a  $p$ -value directly from its distribution.

We present an evaluation of two robust procedures based on the TDT-type association tests. The results appear to be consistent over a variety of sampling mechanisms. When one has little knowledge of the underlying genetic model, our results show that MAX ( $Z_R, Z_A, Z_D$ ) is a more robust test. It provides a better protection against the loss of power under recessive (dominant) models than the TDT or MERT do, while remaining reasonably powerful when additive or multiplicative models hold.

When markers, rather than candidate alleles, are examined the optimal statistics are similar to the ones considered here, however, the correlation structure also depends on the recombination fraction between the marker and putative disease allele. Whether the efficiency robustness approach will yield a single test with high power properties for the four models requires further study.

The MAX test should be extendable to some  $m$ -allele settings discussed, for example, by Sham & Curtis (1995) and reviewed by Ewens (1999), if it is used to replace the TDT in the permutation test approach of Morris *et al.* (1997). When only one of the  $m$ -alleles is related to the disease, this extension of the MAX test should be efficiency robust for the class of models considered in the two-allele setting. However, there are many other possible genetic models in the  $m$ -allele situation, so it is not clear that any single test will be efficiency robust for all of them. Further research is also needed to incorporate other risk factors into the robust tests. In some situations, the age of onset may effect the power of the TDT (Li & Hsu, 2000) and presumably would affect these tests too.

It is a pleasure to thank the referees and Professors R. Curnow, J. Whittaker, A. Morris and W. Ewens for many helpful suggestions that greatly improved the manuscript. The research of J.L.G. was partially supported by a grant from the National Science Foundation.

#### APPENDIX A

First we derive mating type 2, 4 and 5 contributions to the efficient score and information for testing  $H_0: r = 1$  for each of the 4 genetic models using Table 1.

##### Mating type 2

Under the  $D$  model, mating type 2 is not informative and thus it does not contribute to the likelihood. Under  $R$  and  $M$  models the likelihood is,

$$L \sim \left( \frac{r}{r+1} \right)^{n_{22}} \left( \frac{1}{r+1} \right)^{n_{21}}$$

and

$$\frac{\partial \log L}{\partial r} \Big|_{H_0} = \frac{1}{2}(n_{22} - n_{21}), \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{n_2}{4}.$$

Under the  $A$  model

$$L \sim \left( \frac{2r-1}{3r-1} \right)^{n_{22}} \left( \frac{r}{3r-1} \right)^{n_{21}}$$

$$\frac{\partial \log L}{\partial r} \Big|_{H_0} = \frac{1}{2}(n_{22} - n_{21}), \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{n_2}{4}.$$

*Mating type 4*Under the *D* model

$$L \sim \left( \frac{r}{3r+1} \right)^{n_{42}} \left( \frac{2}{3r+1} \right)^{n_{41}} \left( \frac{1}{3r+1} \right)^{n_{40}}$$

and

$$\left. \frac{\partial \log L}{\partial r} \right|_{H_0} = \frac{1}{4}(n_{42} + n_{41} - 3n_{40}), \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{3n_4}{16}.$$

Under the *R* model

$$L \sim \left( \frac{r}{r+3} \right)^{n_{42}} \left( \frac{2}{r+3} \right)^{n_{41}} \left( \frac{1}{r+3} \right)^{n_{40}}$$

and

$$\left. \frac{\partial \log L}{\partial r} \right|_{H_0} = \frac{1}{4}(3n_{42} - n_{41} - n_{40}), \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{3n_4}{16}.$$

Under the *A* model

$$L \sim \left( \frac{2r-1}{4r} \right)^{n_{42}} \left( \frac{2r}{4r} \right)^{n_{41}} \left( \frac{1}{4r} \right)^{n_{40}}$$

and

$$\left. \frac{\partial \log L}{\partial r} \right|_{H_0} = n_{42} - n_{40}, \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{n_4}{2}.$$

Under the *M* model

$$L \sim \left( \frac{r^2}{(r+1)^2} \right)^{n_{42}} \left( \frac{2r}{(r+1)^2} \right)^{n_{41}} \left( \frac{1}{(r+1)^2} \right)^{n_{40}}$$

and

$$\left. \frac{\partial \log L}{\partial r} \right|_{H_0} = n_{42} - n_{40}, \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{n_4}{2}.$$

*Mating type 5*Under the *D*, *A* and *R* models

$$L \sim \left( \frac{r}{r+1} \right)^{n_{51}} \left( \frac{1}{r+1} \right)^{n_{50}}$$

and

$$\left. \frac{\partial \log L}{\partial r} \right|_{H_0} = \frac{1}{2}(n_{51} - n_{50}), \quad I(H_0) = -E \left( \frac{\partial^2 \log L}{\partial r^2} \right) \Big|_{H_0} = \frac{n_5}{4}.$$

Under the *R* model mating type 5 is uninformative and does not contribute to the likelihood.

The numerator and denominator of the efficient score tests (1-3) for each model can now be obtained by summing the corresponding efficient scores and information contributions from each of the 3 mating types.

## APPENDIX B

Suppose the model underlying the data is not known and a family,  $\Psi: \{f_i; i = 1, \dots, I\}$ , of plausible alternative models is specified. Let  $Z_i$  denote the corresponding asymptotically most powerful test for model  $f_i$ . In many situations  $\{Z_i\}$  are asymptotically jointly multivariate normal with correlation

matrix  $\{\rho_{ij}\}$ . The Pitman asymptotic relative efficiency (ARE) of the test  $Z_i$  relative to the test  $Z_j$  when  $Z_j$  is optimal is  $\rho_{ij}^2$ . The correlation matrix,  $\{\rho_{ij}\}$ , of the optimal statistics summarizes the structure of the family of alternative models as each correlation reflects how close, statistically, the two models are.

For any asymptotically normal test statistic  $Z$  denote its relative efficiency to the optimal test  $Z_i$  for model  $f_i$  by  $e(Z, i)$ . The lowest ARE that  $Z$  has when a model in  $\Psi$  is true is denoted  $e(Z, \Psi) = \inf_{1 \leq i \leq I} \{e(Z, i)\}$ . The Maximin Efficiency Robust Test (MERT), satisfies  $e(\text{MERT}, \Psi) = \sup_{Z \in \Gamma} [\inf_{1 \leq i \leq I} \{e(Z, i)\}]$ , where  $\Gamma$  is the set of all consistent asymptotically normal test statistics for the problem. Gastwirth (1966) showed that when the minimum correlation of the optimal test statistics  $Z_i$ ,  $\rho^* = \min(\rho_{ij})$ , is  $> 0$  the MERT exists, is unique, and is a linear combination of the  $\{Z_i\}$ . Another robust test statistic is  $\text{MAX} = \max_{1 \leq i \leq I} (Z_i)$ . Asymptotically, under the null hypothesis,  $\text{MAX}$  is distributed as  $\max[\text{MN}(0, \{\rho_{ij}\})]$ , where MN stands for multivariate normal. Freidlin, Podgor & Gastwirth (1999) showed that when  $\rho^* \leq 0.5$ , the MAX test is more powerful than the MERT, but when  $\rho^* \geq 0.7$ , there was virtually no difference in their powers.

## REFERENCES

- Curnow, R. N., Morris, A. P. & Whittaker, J. C. (1998). Locating genes involved in human diseases. *Applied Statistics* **47**, 63–76.
- Ewens, W. J. & Spielman, R. S. (1995). The transmission disequilibrium test: history, subdivision, and admixture. *Am. J. Hum. Genet.* **57**, 455–464.
- Ewens, W. J. (1999). Statistical aspects of the transmission/disequilibrium test (TDT). In *Statistics in Molecular Biology and Genetics* (ed. Seilier-Moiseiwitsch) IMS.
- Falk, C. T. & Rubinstein, P. (1987). Haplotype relative risk: an easy reliable way to construct a proper control sample for risk calculation. *Ann. Hum. Genet.* **51**, 227–233.
- Field, L. L., Fothergill-Payne, C., Bertrams, J. & Baur, M. P. (1986). HLA-DR effects in a large German IDDM dataset. *Genet. Epidemiol. Suppl.* **1**, 323–328.
- Fleming, T. R. & Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.
- Freidlin, B., Podgor, M. J. & Gastwirth, J. L. (1999). Efficiency robust tests for survival or ordered categorical data. *Biometrics* **55**, 883–886.
- Gastwirth, J. L. (1966). On robust procedures. *Journal of the American Statistical Association* **61**, 929–948.
- Gastwirth, J. L. (1985). The use of maximin efficiency robust tests in combining contingency tables and survival analysis. *Journal of the American Statistical Association* **80**, 380–384.
- Gastwirth, J. L. & Freidlin, B. (2000). On power and efficiency robust linkage tests for affected sibs. *Annals of Human Genetics* **64**, 443–453.
- Lander, E. & Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**, 241–7.
- Li, H. & Hsu, L. (2000). Effects of age at onset on the power of the affected sib pair and transmission/disequilibrium tests. *Ann. Hum. Genet.* **64**, 239–254.
- Morris, A. P., Curnow, R. N. & Whittaker, J. C. (1997). Randomization test of disease-marker associations. *Annals of Human Genetics* **61**, 49–60.
- Ott, J. (1989). Statistical properties of the haplotype relative risk. *Genet. Epidemiol.* **6**, 127–130.
- Rao, C. R. (1973). *Linear Statistical Inference and its Applications*. 2nd ed., Wiley, New York.
- Rubinstein, P., Walker, M., Carpenter, C., Carrier, C., Krassner, J., Falk, C. & Ginsberg, F. (1981). Genetics of HLA disease association: the use haplotype relative risk (HRR) and the 'haplo-delta' (Dh) estimates in juvenile diabetes from three racial groups. *Hum. Immunol.* **3**, 384.
- Schaid, D. J. & Nick, T. G. (1990). Sib-pair linkage tests for disease susceptibility loci: common test vs. the asymptotically most powerful test. *Genet. Epidemiol.* **7**, 359–370.
- Schaid, D. J. & Sommer, S. S. (1993). Genotype relative risks: methods for design and analysis of candidate-gene association studies. *Am. J. Hum. Genet.* **53**, 1114–1126.
- Schaid, D. J. & Sommer, S. S. (1994). Comparison of statistics for candidate-gene association studies using cases and parents. *Am. J. Hum. Genet.* **55**, 402–409.
- Sham, P. (1998). *Statistics in Human Genetics*. Arnold, London.
- Sham, P. C. & Curtis, D. (1995). An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Annals of Human Genetics* **59**, 323–336.
- Shih, M. & Whittemore, A. S. (2001). Allele-sharing among affected relatives: non-parametric methods for identifying genes. *Statistical Methods in Medical Research* **10**, 27–55.
- Slager, S. L., Huang, J. & Vieland, V. J. (2001). Power comparisons between the TDT and two likelihood-based methods. *Gene Epidemiol.* **20**, 192–209.

- Spielman, R. S., McGinnis, R. E. & Ewens, W. J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes melitus (IDDM). *Am. J. Hum. Genet.* **52**, 506-516.
- Spielman, R. S. & Ewens W. J. (1996). The TDT and other family-based tests for linkage disequilibrium and association. *Am. J. Hum. Genet.* **59**, 983-989.
- Tarone, R. E. (1981). On the distribution of the maximum of the log-rank statistic and the modified Wilcoxon statistic. *Biometrics* **37**, 79-85.
- Terwilliger, J. D. & Ott, J. (1992). A haplotype-based 'haplotype relative risk' approach to detecting allelic association. *Hum. Hered.* **42**, 337-346.
- Thomson G., Robinson, W. P., Kuhner, M. K. & Joe, S. (1989). HLA, insulin gene, and Gm associations with IDDM. *Genet. Epidemiol.* **6**, 155-160.
- Whittemore, A. S. & Tu, I. P. (1998). Simple, robust linkage tests for affected sibs. *Am. J. Hum. Genet.* **62**, 1228-1242.